

Early Loan Words in Western Central Asia Indicators of Substrate Populations, Migrations, and Trade Relations

Michael Witzel

Recently discovered evidence suggests that there is a body of loan words, preserved independently from one another in the oldest Indian and Iranian texts, that reflect the pre-Indo-Iranian language(s) spoken in the areas bordering northern Iran and northern Afghanistan, that is, the Bactria – Margiana Complex. These loans include words from agriculture, village and town life, flora and fauna, ritual, and religion. They were taken over and then exported to Iran and northern India by the speakers of the various Old Iranian and Old Indo-Aryan (Vedic) languages as well as a western offshoot, the Mitanni Indo-Aryan of Syria and Iraq and the language of related tribes as indicated by some Indo-Iranian words in Kassite. All these represent series of Indo-Iranian intrusions into the world of the great Mesopotamian, Bactria-Margiana, and Indus civilizations.

INTRODUCTION

Over the past few decades archeologists have discovered an increasing number of sites of the great Oxus Civilization, perhaps better known nowadays as the Bactria-Margiana Archeological Complex (BMAC), as well as its Neolithic and Chalcolithic predecessors.¹ While they have filled in a large gap between the great civilizations of Mesopotamia and the Indus, so far no written documents have been found, with the exception of an anomalous, single inscribed seal from Anau reported by Fredrik Hiebert.² However, little to nothing is known about the languages spoken in the areas east of Mesopotamia (Hurrite, Akkadian, Sumerian, Elamite)³ and those west of the Indus area. The language(s) of the Indus Civilization also are by and large unknown as long as we cannot understand the signs on the Indus seals. Further, the materials that can be distilled from the loan words in the earliest Indo-Aryan texts, the Vedas,⁴ that follow after the collapse of the Indus Civilization, unfortunately have been overlooked for that purpose. They actually can serve as a guide to

what languages may have been present in the subcontinent in around 2000 B.C.E.⁵

The picture can be enlarged and projected back in time by using the oldest Iranian counterparts of the Vedas, the Avestan texts of the Zoroastrians, as well as the Old Persian (O.P.) inscriptions. Since Old Iranian and Old Indo-Aryan (Vedic) are closely related, items common to both languages can be used to reconstruct the common protolanguage, Indo-Iranian.

Apart from this, we have virtually no evidence for the areas between the great civilizations and those north of Greater Iran, as they are too distant from the Near Eastern, Indian, and Chinese cultures to have been discussed or described in detail in their texts.⁶ However, the seal recently discovered at Anau should alert us to the (remote) possibility that early writing might be found in the area after all.⁷ It is unlikely, as the people of the BMAC studiously avoided using writing in their many seals, perhaps as a way of distancing themselves from their Near Eastern neighbors. In the meantime all that we can establish for the languages used in the Western Central Asian area comes from early Near Eastern and Indian (and also Old Iranian) sources. There are some references in the Sumerian and Akkadian documents of the third to first millennium B.C.E., but they deal just with the border areas of Mesopotamia,⁸ and they furnish only vague references, such as that to Aratta, probably Arachosia.⁹ Similarly, we have only a few vague reminiscences in the earliest Indian texts (*R̥gveda*) composed in the Greater Panjab (c. 1200 B.C.E. – 1000 B.C.E.),¹⁰ which seem to refer back to the area along the Volga (Rasā) and, secondly, to the people who dwelt along the River Sindes (Tacitus' name for the Merv or Tedzhen River): the Dāsa or O.P. Daha (whom the Greeks called Da[h]ai), the Arii, and the *Parna (Ved. Paṇi? cf. Ptolemy, Geogr. 6.10.2, Parnoi, Daai; Parni, Dacae, otherwise Dahae). Pinault (2003) connects Paṇi/*Parna, as a loan word from the BMAC area, with Tocharian B peñiyo, A pañi "splendor."

In addition, after the sparse attestation found in the Old Iranian and the much later Middle Iranian sources,¹¹ most of these areas became Turkish-speaking after about 1,400 years ago; this has obliterated much if not most of the older Iranian and Indo-Iranian (IIr.) record, frequently even that of topographical names.¹² Even in this unfortunate situation, we can retrieve, based on the written records of neighboring Indo-Iranian peoples and on old loan words, an increasing number of details of the pre-Indo-Iranian/Iranian languages of the area, notably that of the BMAC (or Oxus Civilization, c. 2400 – 1600 B.C.E.),¹³ and of Greater Afghanistan. However, it is precisely these Indo-Iranian sources that largely have been neglected so far.¹⁴

For some years (1995 – 2001) I have drawn attention, mostly in brief and passing fashion, to a common body of words in Old Indian and Old Iranian

texts that do not seem to be of Proto-Indo-Iranian (nor perforce Proto-Indo-European) origin. These words represent the non-Indo-Iranian languages spoken in Iran and in the northwestern part of the Indian subcontinent at the time these texts were composed, that is, late in the second and early in the first millennium B.C.E. As such, they are invaluable materials for the study of the language(s) preceding the introduction of Indo-Aryan (Vedic) and Old Iranian (Old Persian, Avestan). More important, both types of hieratic texts share a common substratum that can only be that of the southern part of Western Central Asia. As will be seen below, it cannot come from elsewhere: both Vedic and Old Iranian individually imported it into their particular habitats, the Greater Panjab and Iran/Afghanistan.

Such substrate words are quite common in languages that have occupied the territory of an earlier people speaking a different language.¹⁵ In English, for example, such common words as sheep (Dutch *schaap*, German *Schaf*) belong to the Neolithic substratum of the North Sea coast of Northern Germany and Denmark, the homeland of Anglo-Saxon.¹⁶

SOURCES

In order to evaluate properly the scarce materials at our disposal, a brief look at our sources is in order.¹⁷ The Vedas were composed (roughly between 1500–1200 and 500 B.C.E.) in parts of present day Afghanistan, northern Pakistan, and northern India. The oldest text at our disposal is the *R̥gveda* (*RV*); it is composed in archaic Indo-Aryan (Vedic Sanskrit). It is followed by a number of other Vedic texts, usually listed as *Samhitās*, *Brāhmaṇas*, *Āraṇyakas*, and *Upaniṣads*. Linguistically, however, we have to distinguish five distinct levels: *R̥gveda*, other *Samhitās* (Mantra language), *Yajurveda Samhitā* prose, *Brāhmaṇas* (including *Āraṇyakas* and *Upaniṣads*), and the late Vedic *Sūtras* (Witzel 1987, 1997).¹⁸

The language of the *RV* is an archaic form of Indo-European. Its 1,028 hymns are addressed to the gods, and most of them are used in ritual. They were orally composed and strictly preserved until today, by exact repetition through rote learning. It must be underlined that the Vedic texts are “tape recordings” of this archaic period.¹⁹ Not one word, not a syllable, not even a tonal accent was allowed to be changed.²⁰ The texts are therefore better than any manuscript and as good as any well-preserved contemporary inscription. We can rely on the Vedic texts as contemporary sources for names of persons, places, and rivers (Witzel 1999a) and for loan words²¹ from contemporary local languages.²²

The *R̥gveda* was composed in the Greater Panjab²³ and dates from before

the introduction of iron in the northwestern subcontinent. Later texts cover all of northern India up to Bengal and southwards toward the Vindhya hills.

Some four percent of the words in the *Rgvedic* hymns that are composed in an archaic, poetic, hieratic form of Vedic clearly are of non-Indo-European, non-Indo-Aryan origin. In other words, they stem from a pre-Indo-Aryan Panjab substrate.²⁴

The situation is similar but not quite as beneficial where the Old Iranian texts are concerned. Only about a quarter of the original *Avesta* has been preserved. The oldest parts are Zarathustra's *RV*-like poems, his five long *Gāḍās* (*Yasna* 28–53), and his(?) contemporaneous ritual text embedded among the *Gāḍās*, the *Yasna Haptanḥā'iti*, a collection of mantras used for fire worship. The rest of the Avestan texts are post-Zoroastrian and composed in the Young Avestan language. However, the initial oral tradition of the *Avesta* was converted in Sasanidian times (c. 400 C.E.) into a written tradition whose surviving earliest manuscripts are not older than a thousand years and have been corrupted by centuries of decline during the early Islamic period of Iran.²⁵ Nevertheless, the philologically restored Avestan texts offer some data from Greater Afghanistan, because Zarathustra's homeland was probably situated in northwestern Afghanistan (near the Kashaf River),²⁶ and much of the later *Avesta* was composed or redacted in southern Afghanistan (Seistan, Arachosia). However, in spite of being geographically closer to the Mesopotamian cultures with datable historical information, the Avestan texts are even less amenable to absolute dating than the Vedic ones. Mesopotamia (or early China) simply does not figure in all these texts.

The older Avestan texts (*Gāḍās* / *Yasna Haptanḥā'iti*) point to a Copper or Bronze Age culture quite similar to that of the *RV*. The younger texts might overlap with the expansion eastward of the Median realm (c. 700–550 B.C.E.).²⁷ The few Old Persian inscriptions that have survived date from 519 B.C.E. onward. However, unlike those in the old Indian texts,²⁸ the "foreign" words in the Old Iranian texts have not been evaluated so far. Researchers apparently were of the opinion that only a few could be found; the matter simply has been neglected.

LOAN WORDS AND SUBSTRATE LANGUAGES

Words from substrate languages are defined here as all those words in early Vedic that do not conform to Indo-European/Indo-Iranian word structure (including sounds, root structure, and word formation) and that have no clear Indo-European/Indo-Iranian etymology.

We have to distinguish various types of loans (Anttila 1989, 154ff.).

Some are due to cultural and economic contacts, such as the modern "guru" or "karma" (from India), or the slightly older "coffee" (from Arabia), "cocoa," "chocolate" (from Meso-America), or "tea" (French *thé*, and so forth), whose origin can be traced to southern Chinese (Amoy *t'é*), while the Russian, Indian, and Japanese *chai/cha* and Nepali *chiya* are from a northern Chinese dialect. The Indian word *chai* has thus come overland and not by sea. The example would also be instructive if we did not know the history of transmission: linguists would be able to pinpoint the origin of the loan in two areas of East Asia. Similar examples will be found below for Central Asian words. This kind of introduction of loan words is from an "adjoining" language, an adstrate. Examples abound in multilingual societies (India) or societies in close contact (ancient and modern Western Europe, with cases such as street < Latin (*via*) *strata*, tsar/Kaiser < Caesar, castle < *castellum*, cellar < *cellarium*, *cella*, and so on).

Loans stemming from previously existing languages, on introduction of a new, dominant language, are different (Anttila 1989, 1, 171 – 172). The new language may function as superstrate, properly used and understood only by a minority at first (such as Latin in Celtic France), but it then spreads by assimilating a large number of local words from the previous language, the substrate (note the Celtic place-names in England, below).

Sometimes the superstrate does not become dominant (as Norman French failed to do in England); in this case we may still expect a large number of words from the superstrate in the persisting local language (French *beauté* > beauty, *ancêtre* = ancestor, where the English form reveals the older French word, with *-st-*).

If the source of the loan remains unknown, many loan words from "foreign" (substrate/adstrate) languages can still be easily detected by linguistic means, even if they belong to a long disappeared language. The reason is that all languages follow certain patterns, allowing only certain sounds or groups of sounds, while others that are difficult to pronounce must be replaced by locally accepted ones. English, for example, until fairly recently, found German and Yiddish words beginning with the sounds *sh* + consonant (*schnitzel*, *strudel*, *schlep*) impossible to pronounce, as English allowed only *s* + consonant (as in *snit*, *strut*, *slip*). By now, these sounds have been accepted and are pronounced correctly. Similarly, even today words beginning in *ng-*, *mf-*, and so on, are not allowed in English (though by now a few African names have been locally adopted, such as *Mfume*).

Words with such uncharacteristic sounds or sound clusters therefore indicate a certain cultural influence, even if the native speaker (or a latter-day scholar) may not know where these words came from originally.

This is especially true when we have to deal with toponyms and hydronyms that have come down to us from prehistory. It is well known that place-names, especially names of (larger) rivers, are very conservative. Even today they may reflect languages spoken many thousands of years ago. For example, we have the Rhine (Lat. loan word *Rhenus* < Celtic **Rēnos* < IE **reinos*); Danube (Lat. *Danubius* < N. Iran. *Dāna-*); Don (Gr. *Tanais* adapted from **Dāna-*); Tigris (Latin, Greek < O.P. *Tigrā* [cf. O.P. *tigra* "quick"], Arab. [Nahr al] *Dijlat*, both < Akkad. [I]di-iq-lat / Sumerian *Idigna*, all from a pre-Sumerian substrate!); and Euphrates (cf. Arab. [Nahr al] *Furāt*, which has been taken over from Greek < O.P. [h]Ufratu, close to O.P. [h]u-frātar "good brother), Sum. *Buranuna* / Akkad. *Purattum* / Elam. *ú-ip-ra-tu-iš*, again all from a pre-Sumerian substrate of more than five thousand years ago).

The early river names of most of Europe belong to one and the same old system.²⁹ A different prehistoric system is found in Greece and the Aegean area, with the typical pre-Greek *-s(s)-*, *-nt-*, *-mn-* suffixes. This phenomenon was extensively discussed for much of the twentieth century.³⁰ A detailed investigation of both the regions just mentioned can serve as a guide — and a warning — for the following deliberations.

It is important to keep in mind that names taken from a previous language (or from an adstrate) have more often than not lost their (precise) original meaning. If the source language is little known or unknown, we can only analyze and compare the outward form of the names involved. This includes the sound system as well as typical suffixes and prefixes that indicate, as in the case of toponyms, "river, place, mountain, plain" or describe the item in question, such as "quick/slow, white/black" (river), "high/low" (place). However, these names often have been adjusted or reinterpreted by later languages, frequently by popular etymology (see above, Tigris as *tigra* "quick"). As can be seen in the discussion of the pre-Hellenic and Old European place/river names, these conditions can lead to many pitfalls.

The situation of Central Asia can be approached by a comparison with that of place-names in England. We know that the early form of English, an Old Saxon dialect (a part of the Germanic branch of IE), was overlaid, in the middle of the first millennium C.E., on the Celtic (and Latin) languages of Britain. Both Celtic and Latin have left a number of loan words in Old English as substrate words, such as *London* < Celtic *Lugodunum* "town of the god Lug" and *-chester* < Latin *-castrum* "fortified settlement." Later on, English saw the superimposed (superstrate) influences of the Viking language (North Germanic, with words such as "egg" and place-names in *-vik* and *-ay*), then of Norman French, with a large number of loans ("beauty," "ancestor," and so on), and finally an equally huge number of learned, newly formed Greco-Latin

words as well as various minor adstrate influences from neighboring languages such as Dutch ("marine" words such as "dike," "boss," "mate," and so forth). Most interestingly, Old Saxon and Germanic in general can be shown to have a large percentage of non-Indo-European substrate words (such as "sheep," "eel," "roe," "boar," "lentil," "land," "delve," and "prick") derived from a long-lost prehistoric Northern European language).³¹

The situation in the Greater Panjab (the area of the earliest Vedic texts) and in Greater Iran (the area of the Avestan and Old Persian texts) is quite similar. A brief, simplified summary looks like this:

GREATER PANJAB

English loan words
Urdu/Perso-Arabic superstrate loans
influx of learned (Sanskrit) words

Development to New Indo-Aryan
influx of learned (Sanskrit) words

Development of various Prākṛts
Old Greek loan words
Old Persian / Iranian loan words

Development from Vedic to Middle Indo-Aryan

Later (R̥g) Vedic / OIA dialects
Dravidian adstrate
immigrant OIA

Harappan language (see text below)

unknown local language(s)

GREATER IRAN

Development to modern Iranian
Arabic superstrate loans

Development to Middle Iranian
some Old Greek loans
Old Persian / Later Avestan

Old Avestan
Old Iranian superstrate
immigrant OIA in Iran
(*Sarasvatī, Sarayu, etc.)
Central Asian substrate
Indo-Iranian in Central Asia,
south of Uralic, Ket
(Yeneseian)
immigrant Indo-European

DELINEATING THE BACTRIA-MARGIANA ARCHEOLOGICAL COMPLEX

All of the types of data mentioned so far need to be studied in greater detail, especially the reconstruction of an early Indo-Iranian presence in Central Asia

and on the Iranian Plateau based on linguistic, religious, social, and material culture – related data (Witzel 1999b, 1999c). As mentioned, for the BMAC area itself, we have no written sources at all, except for the loans quoted above (for further details see below).

First it has to be established why one should think of Central Asian origins for the loans found in both Indo-Aryan and Iranian. In theory, such loans could also have originated in the Panjab and have traveled to Iran (as, indeed, the word for the new crop, “rice,” did); or, conversely, they could have traveled from an Iranian area to the Panjab (as is the case with the Near Eastern word for “wheat”).

However, words that can be reasonably well plotted in both time and place, that is, **uštr* “camel,” **khar* “donkey,” and **išt* “brick,” point to the areas along the northern rims of Greater Iran (BMAC for short). This is the first Central Asian area with a highly-developed agriculture and town civilization with which the speakers of Indo-Iranian could have come into contact, south of their more original habitat, exemplified by the contact zone with the Uralic and Yeneseian (Ket) speaking peoples.³² We know that, in the BMAC area, the domesticated camel was used,³³ the large-scale use of unburned bricks was continued, and the donkey was introduced from the Near East at the time. These three leitfossils also provide a time frame: the speakers of Indo-Iranian cannot have moved through this complex earlier than the introduction of donkeys.³⁴ It is important to note that Pinault (2003) has now also provided the eastern rim of influence of the BMAC language(s) by identifying some words that have been taken over into early (Common) Tocharian, such as *išt*, *anc'u*, *c'arwa*, *pani*, *āñi*, and *athr* (see below).

However, as will be seen, some of the Indo-Iranian loan words have been taken over twice or thrice, independently of one another, into various Indo-Iranian languages in different areas of Greater Iran, but such words still look very similar to each other. This kind of difference is a clear indication of dialect variations in an underlying substrate language. A selection of the most typical loans tends to center on the Bactria-Margiana area (see Witzel 1995, 1999b, 1999c, 2000) with perhaps the inclusion for some words of Seistan/Arachosia, where the delineation of the boundary of the “BMAC” language is unclear so far.

However, for other parts of Greater Iran (modern Iran proper, Baluchistan, Afghanistan), the southern boundary of the “BMAC” language(s) is fairly clear. We can easily exclude the southern belt of Iran as well as Baluchistan. East of Akkadian and Sumerian in Mesopotamia, the isolated language Elamite was spoken,³⁵ not just in the Susiana (Khuzistan) and Anšan (Fars and surroundings), but also in southern and central Iran,³⁶ in the areas

of Tepe Yahya (Simaški)³⁷ and Shahdad (Tukriš),³⁸ but apparently not farther east than Bampur (Marhaši).³⁹

According to F. Vallat, the language of Marhaši⁴⁰ differs from that of Simaški, and it is only minimally Elamite-related (1985, 52). Here and in Meluhha (Baluchistan, Sindh)⁴¹ the language changes, though Elamite cultural influence extends even to the borders of Sindh. However, eastern names are neither Elamite nor Dravidian.⁴² This indicates that there was a language boundary somewhere to the west of the present Iran-Pakistan border, probably in a southward prolongation of the Iran-Afghanistan border. That the Meluhhan language was different altogether from Elamite or Sumerian is obvious: the Mesopotamians needed a "translator from Meluhha" (Possehl 1996, no. 2), whose name was reported as Šu-ilišu (Parpola 1994, 132).⁴³

Baluchistan is characterized by an overlap between the influences from the Elamite and Indus cultures. Baluchi sites such as Kulli are dated about 2000 B.C.E.⁴⁴ The area that was called in Old Persian Maka (Mesopot. Makkan), or now Makrān, has its center at Marhaši (Bampur), an important place of exchange between the Indus areas to the east, the Elamite areas to the west, the Arachosian site of Mundigak, and Shahr-i-Sokhta (Aratta, Seistan), Simaški, as well as Turkmenia to the north. Only during the last period of Mehrgarh (level VII) do we find a cemetery with BMAC-derived items, and here the rituals and metallurgy recall those of the BMAC as well. The spread of late BMAC influences into the Indus Valley, Baluchistan, Susiana, and so on, is now well documented.

While the southern belt of Greater Iran is thus excluded, the exact alignment of the BMAC language boundary in the more northern parts of Greater Iran is not so clear. We have some indication that a series of peoples entered Mesopotamia from the Zagros Mountains, starting around the end of the third millennium. First came the Lullubi and the Gutí, in the time of Narām Sīn of Akkade (2213–2176 B.C.E.). On the Iranian highlands, the Simaški state of the third millennium is replaced by the Tukriš state in the second millennium, which is connected, in some texts, with the Gutí.⁴⁵ The Hurrites (with a Caucasian language related to the later Urartian) appeared in the same epoch but more to the north; the Kassites entered under the successor of Hammurabi, Šamšu-Iluna, at c. 1740 B.C.E. and soon took over Babylon for centuries. Both the Kassite (Balkan 1954) and Hurrite languages (Mitanni, around 1400 B.C.E.) have some Old Indo-Aryan linguistic elements.⁴⁶ However, we know much less about Media (the area around Tehran), Herat, Seistan, and Arachosia.

While the language of the substrate studied below included the greater BMAC area, roughly from Anau to Balkh, we cannot firmly exclude, at this

stage, the Media/Areia/Arachosia areas as belonging to that of the BMAC language.⁴⁷

Indeed, **anc'u*, “soma plant” (probably ephedra, discussion below), which is pressed to prepare the sacred drink of the Indo-Iranian peoples, points to the high mountains of Central Asia (including the Hindukush, the Pamirs, and the Himalayas; see Staal 2001), where, according to both the *Avesta* and the *R̥gveda*, the best soma grows. There are, indeed, some indications of non-Indo-Iranian speakers in the high mountains of Afghanistan even at the time of the earliest Indo-Iranian texts. Such “foreigners” indeed still survive in the Pamirs,⁴⁸ with Burushaski speakers, and are visible in the substrate in Khowar (Witzel 1999b, 1999c, 2000).

AN OLDER CENTRAL ASIAN LEVEL: THE OXUS/BMAC LANGUAGE(S)

At this stage it is worthwhile to take a close look at the actual linguistic evidence preserved by the loans into the Indo-Iranian languages. Such Indo-Iranian words do not have Indo-European shape and etymologies (Witzel 1995, 1999b, 1999c; see also Lubotsky 2001).⁴⁹

Shibboleths for time and place are, as mentioned, the local words **uštr* “camel,” **khar* “donkey,” and **išt* “brick.”

The Bactrian camel was domesticated in Central Asia in the late third millennium B.C.E. and introduced in the BMAC area late in the third millennium, around 2000 B.C.E. (Meadow 1983; Masson 1992, 39–40, 229, 233). It is also found on a few Indus copper plates. Its Mesopotamian designation, found in Middle and New Akkadian *udru* “Bactrian camel,” is a loan from Iran (Mayrhofer 1986–1996 [hereafter *EWA*], 1: 238, Mayrhofer 1956–1976 [hereafter *KEWA*], 3: 652, cf. Diakonoff in *JAOS* 105 [1985]: 600). However, the Indo-Iranian designation, of Proto-Indo-Iranian **uštra*, represented by the identical Vedic *uṣṭra*, Avestan *uštra* “camel,” has no plausible Indo-Iranian or Indo-European etymology.

The case of the donkey is similar, though the source of the word seems to be a southern one. Indo-Iranian **khara* > Vedic *khara*, Avestan *xara* “donkey” seem to go back to the earlier areas of its domestication, in this case Mesopotamia, where an Akkadian (Mari) source has *h̄arum*, *ajarum* “male donkey” (*EWA* 1: 447, 473). Various types of donkeys can be distinguished in Vedic India,⁵⁰ in addition to the other wild equid, the hemione (onager, *Equus hemionus khur*). Interestingly, Vedic, Tocharian, and Dravidian have words that resemble the Indo-Iranian and Akkadian form: Ved. *garda-bha* “donkey” :: **karcabha?* > Toch. B *kercapo*⁵¹ :: PDrav. **garda(?)* > Tamil *kalutai* “donkey.”⁵² They

all may go back to a Near Eastern form **xar-* that is also reflected by Indo-Iranian **khara*. When and how Dravidian took over the word remains unclear so far (cf., however, Blažek 1992, 2002).

Finally, the word for “brick” finds its likely source right in the center of these relationships, in the BMAC area. An Indo-Iranian or Indo-European etymology is unlikely (cf. *EWA* 1: 201; Lubotsky 2001, 311) for Ved. *iṣṭi*, *iṣṭakā*, *iṣṭikā*; Avest. *iṣṭiia* “brick,” *zamō-iṣṭuua* “clay brick”; O.P. *iṣṭi*; M.P., N.P. *xiṣṭ*. One also can compare Toch. B *iṣcem*, *iṣcake* “clay” (i-stem, Toch. B **iṣce*, Toch. A **iśāc* < BMAC **iṣṭ[i]*; Pinault 2003), and, unnoticed so far, Bur. *d-iṣṭik*, Shina *d-iṣṭik*, Kalash *kh-iṣṭi-poktā*, Marathi *v-iṭ*, “brick” (instead of expected **iṭ*). All of this leads to a Central Asian source for “(clay) brick,” **iṣṭ*, with unclear pre-Indic “prefixes” *d-*, *v-*.

Since there are no brick buildings at the northern rim of Central Asia,⁵³ and as both the Indus and Mesopotamian areas are too asymmetrically located to be the source of all these words, the most likely Central Asian source for all these loans is the BMAC area. Instead of some hypothetical Caucasian source, the BMAC area is the first place where the steppe people, the speakers of Indo-Iranian, coming from somewhere north, close to Uralic and Yeneseian (Ket), would first encounter bricks and would need a term for them. As discussed above, the domesticated camel was used here, and the donkey had been introduced from the Near East at the time. These three leitfossils provide the time frame.⁵⁴ This is confirmed by the forms of some words in still undivided Indo-Iranian, which can be dated at c. 2000 B.C.E.: the common Indo-Iranian designations for the newly invented horse-drawn chariot (**ratha*, Ved. *ratha*, Ir. *raṣa*, Mitanni *Tuṣ-ratta*, Kassite *Abi-rattaš*) and related words (**rathin*, *rathī/raṣī* “chariot driver,” and the archaic compound words *rathēṣṭhā/raṣaēṣṭa* “chariot warrior”).⁵⁵ All of this leads to a successful triangulation of material culture, time, and space in Western Central Asia of c. 2000 B.C.E. or slightly later.

WHEAT AGRICULTURE

To these examples, the plants cultivated in the area can be added. Wheat probably is the best case. It is a western import, as it originated west of the Zagros and south of the Caucasus,⁵⁶ in the western Fertile Crescent (between Jericho and Jerf al Ahmar). In South Asia it is found as early as the seventh millennium B.C.E.; it is first attested linguistically by Ved. *godhūma*, Drav. (Kan.) *gōdi*, Bur. *guriñ*, *gureñ* (pl.), *γòrum*, all derived from West Asia.

Ved. *godhūma*, Avest. *gañtuma* must go back, with folk etymology on the Indian side (see below),⁵⁷ to a common IIr. < **gant-um*,⁵⁸ which, in turn, echoes

a Near Eastern source **qend*: Semit. **hnt* (Arab. *hint atum*), Hitt. *kant*, Egypt. *xnd* (*EWA* 2: 499). In contrast, a Caucasian/Bur./Basque (Macro-Caucasian) source is seen in Bur. *gur* “wheat”⁵⁹ ~ Basque *gari* “wheat,” Proto-Kartvelian (Georgian, etc.) **ghomu*, cf. Proto – East Caucasian (Daghestani, etc.) **Gōl’e* (see Witzel 1999a, 1999b, 1999c). The ultimate Near Eastern source for all these words must have been something like **gar* // *q/gan-d* (ə = uncertain vowel).⁶⁰

The Indo-Iranian source **gant + um-a* differs from its Near Eastern source **gand-* by a suffix *-um*, which is attested not only in Iranian (Avest. *gant-uma*⁶¹ and Vedic *godh-ūma*⁶² but also in Burushaski **yond-um* (and perhaps in PKartv. **ghomu* “millet”). One can assume a word **gant-um* that has entered Iran and India via the northern Iranian trade route (Media-Turkmenistan-Margiana / Bactria-Aratta / Seistan), while the forms with *go-* (PKartv. **ghomu*, PEC. **Gōl’e*), which are reflected in Dravidian, must have come via the southern route (Elam/Anšan – Simaški/Tepe Yahya – Marhaši/Bampūr), resulting in Drav. **gōdi* (Kan. *gōdi*, Tam. *kōti*; cf. Burrows and Emeneau 1984, 1906).⁶³ The post-*RV* change from *-an-* > *-o-* is untypical for the Panjab but common in Sindh (Witzel 1999b, 1999c). Secondarily, it must have influenced, in some way (such as Dravidian influence on the late *RV*; Witzel 1999c), the actual form of Ved. *go-dhūma* as well. Note that the Dravidian word, too, seems to be based on a popular etymology.⁶⁴

Interestingly, irrigation channels, the only means of sustainable cultivation in these dry lands, are indicated by a loan word as well, **ya(u)vyā* “streamlet, channel.”⁶⁵ Irrigation agriculture is said to have existed in Western Central Asia from 2200 B.C.E.⁶⁶ The difference in sound between Vedic and Old Persian indicates a slightly different source of the loan, for example, more eastern rather than a western or Gorgan origin: Ved. *yavyā* “stream, channel,” but O.P. *yauviyā* > M.P., N.P. *jō, jōy* “stream, channel,” Parachi *žī* “rivulet” (*EWA* 2: 405). Both forms cannot go back to exactly the same source; this kind of telling difference is often indicative of loan words.⁶⁷ In the next section, a few more important words relating to material culture are listed briefly.

OTHER TERMS OF AGRICULTURE AND VILLAGE LIFE

Items from agriculture and settled life include words for “sheaf,” “seed, semen,” “pillar,” “to heal, healer,” “lute,” “mark,” “lump,” and even colors: “blue,” “brown.” These items are indicative of the expanded agricultural base and the settlements of the BMAC. They are represented here without the typical Indo-Iranian stem vowel *-a-*; we do not know the original stem forms or word-final sounds of the BMAC language(s).

- **parś*, “sheaf”: Ved. *parśa* “sheaf (of corn),” Y. Avest. *parša* “sheaf” (EWA 2: 101)
- **bīj* “seed, semen”: Ved. *bija* “seed, semen,” OIran. **bīza* (in names), Buddh. Sogd. *byz’k*, Parachi *bīz* “grains”
- **sthūnā* “pillar”: Ved. *sthūna*, Y. Avest. *stūnā*, *stunā*, O.P. *stūnā* “pillar”; unless ~ Ved. *sthūra* “tall, thick,” Avest. *-stura*, Khot. *stura* (thus EWA 2: 768)
- **pind* “lump”: Ved. *piṇḍa* “lump, ball (of food),” Khotan. *piṇḍaa*, Armen. *pind* “compact, firm” < Iran.⁶⁸
- **ling* “mark”: Ved. *linga* “mark, penis”: Avest. *haptō-iringa* “the seven marks” = the seven stars of the Great Bear / Wain (*ursa maior*) (see EWA 2: 478–479)
- **vinā* “lute”: *vinā* “lute” (YV+ RVKh), Khot. *bīna* “harp, lute,” Sogd. *wyn’* “lute,” M.P. *win* “lute,” Armen. *vin* “lute” — unless these are loans from India (cf. EWA 2: 568)
- **mani* “jewel,” Mitanni IA *manni-nnu*, Ved. *maṇi*, Avest. *-ma’ni* — if not from IE; cf. Lat. *monile*, OEng. *mene*, and so forth (see EWA 2: 293–294)
- sīs* “lead” (EWA 2: 734), from southwest Iran; (O.P.) **siša* “white,” which would be too early for a loan into Old Persian as *sīsa*,⁶⁹ is found already in AV — rather the word must be due to a local substrate,⁷⁰ perhaps one found in the Kandahar area⁷¹
- **āni* “lynch pin,” Ved. *āṇi*; Pinault 2003 connects Toch. B *oñiye*, loc. *oñi-ne* < **āni-ēn* “hip” (*pars pro toto*, of a vehicle)
- **kapaut* “blue”: Ved. *kapota* “pigeon”, O.P. *kapauta(ka)* “blue, lapis lazuli”; Khot. *kavūta* “blue”, M.P. *kabōd* “gray-blue,” *kabōtar* “pigeon” (see EWA 1: 303; Kuiper 1991)
- **kadru* “brown”: Ved. *kadru* “red-brown,” *Kadrū* “a snake deity,” Avest. *kadruua.aspa* “with brown horses,” N.P. *kahar* “light brown”

LOCAL PLANTS

Importantly, some Central Asian plants are included in this list as well.

- **bhang* “hemp”: Ved. *bhaṅga*, Iran. *banga* “hemp, hashish,” if not ~ *bhañj* “to break,” cf. Iran. *Hauma-varga* “soma twisters,” a Saka tribe; cf., however, Macro-Caucasian words such as Bur. *bañ*, *bañgi* “hemp,” Caucasian: Chechen *baga* “pine tree,” Lezgi *muk’-rag* “fir tree,” etc. < PEC
- **bhink’wV* (Witzel 1999b; Bengtson 2001)
- **anc’u* “soma/haoma plant”; see discussion above.

The following two cases, however, indicate a much deeper level and a wider spread, extending, via North Iranian languages (such as Scythian), to the borders of Europe. These *Wanderwörter*, words of culture such as “coffee” or “tea,” have traveled the length and breadth of Eurasia.

****kan-/k'an-** “hemp”: Ved. *śaṇa* “hemp, cannabis,” M.P. *šan* “hemp,” Khot. *kaṃha*, Osset. *gæn*, *gæncæ*; note: Sumer. *kunibu*; Gr. *kánnabis*, Russ. Church Sl. *konoplja*, OHG *hennep* < Scythian **kanap*; Kirgiz, etc., *kändir*, Turkish *kendir* > Hung. *kender*; Mordwin. *kañc't*, *kañf*; Cheremis *kəñe* — for the change between *k'/ś* in E. Afghanistan/Panjab as in *Karkōṭa/Śarkoṭa*, see Witzel 1999b, 1999c

****sinšap** “mustard”: Ved. *sarṣapa* “mustard” (*Brāhmaṇas*) > MIA, NIA *sāsapa* “mustard seed,” Khot. *śśaśvāna*, Parth. *šyṣ-d'n*, Sogd. *šywšp-δn*, M.P. *span-dān*, N.P. *sipan-dān* “mustard seed” (cf. Gr. *sīnāpi*) < pre-Iran. **sinšapa* < ****sinsap** (Henning's *s₁ens₂ap*; see KEWA.)⁷²

LOCAL ANIMALS

Even more important, some names for local animals pinpoint the area of the “BMAC” substrate more clearly. One would imagine that the Indo-Iranians had a word for the common tortoise. However, this word is a loan that has been used for a river as well.

***kac'yap** “tortoise”: Ved. *kaśyapa* KS+, Y.Avest. *kasiiapa*, Sogd. *kyšph*, N.P. *kašaf*, *kaš (a)p* “tortoise”; cf. Kashaf Rūd, river on the border of Turkmenistan and Khorasan (cf. Humbach 1984, 1991)

The leopard, a common Central Asian animal, later used for sport (see chapter 4 in this volume), again has designations that differ considerably in the various languages concerned. It is already seen as a prominent motif on the oldest pottery from W. Turkmenistan (Dani and Masson 1992).

****pard/pandh** “spotted animal, panther”: Ved. *prḍāku* “snake” RV, *prḍāku* “panther snake” AV, *prḍākhu* BŚS, W.Panj. *parṛā*, O.Iran. **pard-*, cf. Khovar *purdūm* < **prḍhūma?*, Bur. (Yasin) *phúrdum* “adder, snake” (see EWA 2: 163; KEWA 2: 335; Turner 1966, 8362)

The overlap between “panther” and “(spotted) snake” exists because snakes are frequently named after a number of characteristics of other animals

and plants, in this case, the spots of the panther (cf. *EWA* 2: 163). Lubotsky (2nd Intl. Vedic Workshop, Kyoto 1999) has pointed out that the *AV* snake *pr̥dāku* is named after the leopard.⁷³

The Iranian forms differ considerably: N.P. *palang* “leopard, *Felis pardalis*” < O.Iran. **pard-*,⁷⁴ Kurd. *pilink*, Pashto *prang*, Parachi *parō:n* (cf. NIA: Lahnda *parrā*), cf. also Gr. *párdalis*, *párdos*, *léo-pardos*, “leopard,” all < **pard* “spotted, wild animal?” (cf. *EWA* 2: 133), or, following Henning, from an older ***parǵ*. However, the alternate Greek word *pánthēr* must be taken into consideration as well. The original Central Asian word seems to have had the dialect variants ***pard/pand*.⁷⁵

Even more intriguing is the case of the word for the rhinoceros (Witzel 1995, 1999a, 1999b). It seems to go back to a local word **kart/kard* with a “suffix” *-ka/ga* (cf. Lubotsky 2001, on the common BMAC suffix *-ka* added directly to roots).

***kart-ka* “rhinoceros”: Ved. *khadga* “rhinoceros” *MS+*; cf. N.P. *karka-dān*, Arab. *karkaddan*, Aelianus *kartázōnos* (**kargazōnos*), “Indian rhinoceros” (however, cf. Kuiper 1948, 136ff.).

The rhinoceros is by definition a South Asian (Indus Civilization, for example), not a Central Asian, animal. However, its prehistoric spread still needs to be investigated. The local climate and water table around the Seistan lakes and the reed thickets of the Helmand River would have allowed for its existence in the area,⁷⁶ while the comparable situation in the Central Asian oases and their rivers, such as the Balkh River and the Amu Darya, must be studied more closely to determine whether this is a loan from South Asia or a local word as well.⁷⁷ At any rate, based on this word alone, the southern and southeastern limits of the BMAC language cannot be decided, although it offers a welcome possibility of checking the origin, track, and spread of such loans.

THE RELIGIOUS SPHERE

In addition to and beyond the items of material culture listed above, the religious sphere,⁷⁸ too, is strongly involved in the Central Asian loans into Old Iranian and Old Indo-Aryan.

The most prominent words are those of rituals, deities, and priests: **ancu* “soma plant”; **yātu* “black magic,” **atharwan* “priest” (however, see *EWA* 1: 60), **ṛši* “seer,” **uc’ig* “sacrificing priest,” **c’arwa* “name of Rudra,” **indra*, **g(h)andharw/b(h)a* “demi-god or demon.”⁷⁹

All these words are at the center of much of Vedic and also (pre-)Zoroastrian religion, but they have not been considered non-Indo-Iranian (non-IE) so far. The advances made in the study of the Central Asian substrate (Witzel 1995 and later; Lubotsky 2001) now allow us to place these items in perspective.

First, the older Indo-European ritual stressed the ritual drink made from honey, Greek ambrosia (the drink of "immortality" < IE **h₂-m₂rto-*); it is called *madhu*, "mead, fermented honey," in India (see above, on "honey"). Reminiscences of this drink remain in the poetry used for Vedic ritual, where mead has been substituted altogether by the new "pressed out" (*su-*) drink, soma, whose preparation and use developed into the most important Indo-Iranian ritual.⁸⁰ The soma plant, whose botanical nature is still uncertain,⁸¹ originally seems to have been called **anc'u* (Ved. *aṃśu*, Avest. *qsu*; see Thompson 2001).⁸² As has been mentioned, the best variety grows, according to Avestan and Vedic sources, on the high mountains (Tian Shan, Pamir, Himalayas),⁸³ and that is exactly where the more potent variety of ephedra is found. From there it is traded and brought in by the local (mountain) people.⁸⁴ There are indications, so far largely neglected, of a Central Asian ritual involving smaller and larger (soma?) vessels, usually found turned upside down in archeological sites.⁸⁵

Not unexpectedly then, the names of some of the most important Indo-Iranian priests and composers of ritual poetry also belong to this substrate: **atharwan* "(fire?) priest," **ṛṣi*, "seer, poet," **uc'ig* "sacrificing priest," **yātu* "black magic."⁸⁶ Indo-European etymologies, however, not very convincing ones, have been proposed for some of these designations, especially in the light of the analysis given below (*-arwa* suffix).

If these points are evaluated against a discussion of the names of deities, it becomes obvious that a major change in ritual and religion took place among the immigrating speakers of Indo-Iranian in Central Asia. Another feature of Indo-Iranian religion is the emergence of a group of "gods of law and order" (*asura*, *Āditya*) that regulate the "truthful" behavior of people, from a single person to a tribe.⁸⁷ However, other important gods that seem to be based on a substrate designation include **c'arwa* "name of Rudra", **indra*, **g(h)andharw/b(h)a* "a demigod or demon."⁸⁸ These are, however, of uncertain attribution, as they are, outside Vedic India, found only in a late Avestan text (*Vīdēvdād*) and may represent loans from Old Indo-Aryan, when Zoroastrian religion was confronted with Indian "unbelievers."

Importantly, Lubotsky (2001) draws attention to the common "suffix" **-arwa-*⁸⁹ in **atharwan*, *g(h)andharw/b(h)a*, and perhaps also **c'arwa*.⁹⁰ Interestingly, this is seen only in religious terms, which at a minimum points to intentional, if not artificial, formations involving these words.⁹¹

Finally, there are a few important words relating to healers/shamans: **bhiš* “to heal,” *bhiš-aj* “healer”: IIr. **bhiš-aj* > Ved. *bhiṣ-aj*, *bheṣaja* ~ Avest. *baēsaž-iiā*; note the unusual noun formation (*EWA* 2: 264). Shamanism is seen in the BMAC area in a seal depicting a procession with typical flat shamanic drums (Francfort 1994).

The surprising cluster of words related to religion indicates that the Indo-Iranians were strongly influenced by a southern Central Asian population having its own peculiar religion unattested in texts; this whole problem is in need of a close evaluation in comparison with Vedic and Avestan and the remnants of Hindukush (Nuristani, Kalash) religion.⁹²

CONCLUSION

Although we cannot say for sure who the “BMAC people” and their immediate neighbors and predecessors were, by studying their linguistic residue in Indic and Iranian languages, it is possible to learn a great deal about their society and culture. For such a shadowy people — with no writing — it is remarkable that the analysis of loan words reveals concrete data about their building techniques, agriculture, animal domestication, religion, ritual, and other aspects of daily and spiritual life. Much of this analysis can be corroborated by future comparisons of BMAC archeology, involving works such as Francfort’s study (1994) of the religion underlying the BMAC seals and statues.⁹³

In order to recover still greater understanding of this lost Central Asian civilization, it will be necessary to carry out further linguistic investigations — such as a study of the non-Indo-Iranian loans found only in Iranian — and then compare the results with the archeological record. Cooperation with geneticists, physical anthropologists, and other specialists will also yield valuable clues to the identity of the people(s) that inhabited the area of the BMAC.

One thing is certain, however, and that is that the inhabitants of the Bactria-Margiana Archeological Complex, though they lived in a seemingly remote, harsh environment, were not at all cut off from the rest of the world. By all the evidence available to us, it is clear that they interacted, in all four directions, with their neighbors and indeed with more distant communities such as the great West and South Asian civilizations. In short, they behaved like humans everywhere else.

ABBREVIATIONS

Akkad.	Akkadian
Arab.	Arabic
Armen.	Armenian
Austro-As.	Austro-Asiatic
<i>AV</i>	<i>Atharvaveda Samhitā</i>
<i>AV+</i>	Found in <i>AV</i> and later texts
Avest.	Avestan
BMAC	Bactria-Margiana Archeological Complex
<i>BŚS</i>	<i>Baudhāyana Śrautasūtra</i>
Bur.	Burushaski
Cauc.	Caucasian
Drav.	Dravidian
Elam.	Elamite
<i>EWA</i>	Mayrhofer, <i>Etymologisches Wörterbuch des Altindoarischen</i>
FU	Finno-Ugrian
F-Volg.	Finnish-Volgaic
Gr.	Greek
Hitt.	Hittite
IA	Indo-Aryan
IE	Indo-European
IIr.	Indo-Iranian
Indo-Ar.	Indo-Aryan
Ir(an).	Iranian
<i>JAOS</i>	<i>Journal of the American Oriental Society</i>
Kan.	Kannada, Canarese
Kartv.	Kartvelian
<i>KEWA</i>	Mayrhofer, <i>Kurzgefasstes etymologisches Wörterbuch des Altindischen</i>
Khot.	Khotanese Saka
<i>KS+</i>	found in <i>Kaṭha Samhitā</i> and later texts
Lat.	Latin
M	Middle
MIA	Middle Indo-Aryan
Mit.	Mitanni
M.P.	Middle Persian
<i>MS</i>	<i>Maitrāyaṇi Samhitā</i>
N	New
NEC	Northeast Caucasian
Nep.	Nepali
N.P.	New Persian
NIA	New Indo-Aryan
Nur.	Nuristani (Kafiri)
O	Old
O.Avest	Old Avestan
OHG	Old High German

OIA	Old Indo-Aryan
O.P.	Old Persian
Osset.	Ossetic
P	Proto-
Panj.	Panjabi
PEC	Proto-East Caucasian
Pkt.	Prakrit
<i>RV</i>	<i>Ṛgveda Samhitā</i>
<i>RV+</i>	found in <i>RV</i> and later texts
<i>RVKh</i>	<i>Ṛgveda Khila</i>
Skt.	Sanskrit
Sogd.	Sogdian
Sum(er).	Sumerian
Tam.	Tamil
Toch.	Tocharian
Ved.	Vedic
Y.Avest.	Young Avestan
<i>YV</i>	<i>Yajurveda (-Samhitā)</i>

NOTES

A fuller version of this chapter can be found in the author's "Linguistic Evidence for Cultural Exchange in Prehistoric Western Central Asia," *Sino-Platonic Papers* 129 (December 2003). That article contains detailed information on the languages surrounding the Central Asian area, a discussion of place-names and hydronymy, and other related topics.

1. Sarianidi 1992, 1998; Dani and Masson 1992; Francfort 1989, 1990, 1994, 1999; Francfort and Kuz'mina 1998; Hiebert 1998, 2001.

2. Hiebert 2001, 2002; Colarusso 2002; Mair 2001. A few Elamite seals have been found in southern Turkmenistan.

3. Languages known from barely more than the names given to their speakers in Mesopotamian sources, such as those of the Gutti and Lullubi at 2300 B.C.E., are neglected here. For the contemporary situation on the Iranian Plateau, see Vallat 1980, 1985; Steinkeller 1982, 1989.

4. For such (loan) words see Witzel 1999b, 1999c.

5. For a detailed discussion and evaluation, see Witzel 1999b, 1999c, 2001, and forthc.

6. For some such data see below, where the linguistic boundaries of Western Central Asia are discussed.

7. See, however, Mair 2001, yet note Proto-Elamite seals close by at Tepe Hissar and recent finds to the west of Tehran at Tepe Uzbeki (as well as older finds at Shahdad, Shahr-i Sokhta, and so on).

8. Such as the Gutti and the Lullubi; similarly, the texts of the Hittite and Urartu realms for the boundary areas of Anatolia, northwestern Iran.

9. Connected with lapis lazuli, which is found in the nearby Chagai Hills (just south

of Arachosia) and, more famously, in Badakhshan. Note Steinkeller 1982, 250, with details about a green variety, "carnelian with green spots," possibly turquoise, from Marḥaši. However, the blue Badakhshan variety has been most famous. For Aratta see Steinkeller 1982; Vasil'kov and Gurov 1995.

10. The earliest date depends on the date of iron, c. 1000 B.C.E.; see Possehl and Gullapalli 1999. For present purposes "Greater Panjab" indicates the area from Gandhāra (Peshawar) and Swat in the west to Delhi and the Upper Doāb in the east, from the lower Pamir/Himalayan ranges in the north to the borders of Sindh and the Bolan in the south; however, the clear center of the Rgvedic area is western and eastern Panjab/Haryana.

11. A few texts in Bactrian, Khorezmian, and Parthian as well as (frequently mythical) data in Pahlavi and in the Greco-Roman sources; for the eastern areas we have a number of Sogdian texts.

12. Note the proliferation of Central Asian place-names ending in Turkic *-su*, *-kul*, *-kum*, and so on. See the essay by Peter Golden in this volume for information concerning the rise and spread of the Turks. Needless to say we hardly have adequate etymological dictionaries for the Turkic or even for the Iranian languages (with the exception of one for older Turkic by G. Clauson, the dated one of P. Horn [1893] for Persian, G. Morgenstierne [1927] for Pashto, and H. Bailey for Khotanese Saka).

13. Based on new carbon dates; see Francfort and Kuz'mina 1998, 468. "Greater Afghanistan" signifies the territory covered by this country and some adjacent surrounding areas.

14. For initial suggestions see Witzel 1995, introduction to Kuiper 1997, xix; Witzel 1999b, 1999c, 2001; see further Lubotsky 2001.

15. Exceptions are the territories of Australia, Polynesia, and the Americas when first settled.

16. Cf. Huld 1990; Polomé 1986, 1990; Vennemann 1994, 1998.

17. For the sparse Mesopotamian sources, see below (Steinkeller, Vallat).

18. For abbreviations of the names of texts, their dates, and their geographic locations, see Witzel 1987, 1989, 1997.

19. The middle and late Vedic redaction of the texts influenced only a very small, well-known number of cases, such as the development *Cuv* > *Cv*.

20. They even preserve very special cases of sentence intonation; see Klein 1997; Witzel 2001.

21. Summary and discussion for *RV* words by Kuiper 1991; for post-*RV* texts, see Witzel 1999b, c.

22. The Vedas are followed by the ancient Tamil "Sangam" (*Caṅkam*) texts from the beginning of our era, all virtually unexplored for substrates and adstrates. On the Iranian side, there are sources such as the Pahlavi and early New Persian texts (*Šah Nameh*, and so on), all beyond the scope of the present essay. For Iranian place-names, see Eilers 1982, 1987; Savina 1964; Schmitt 1995. Such investigations, however, are largely lacking for Afghanistan (note, however, Gryunberg 1980; Pakhalina 1976; Rozenfel'd 1953 for the extreme northeast).

23. See Witzel 1997, 2001: roughly from eastern Afghanistan, Gandhara, Panjab, up to Delhi and even up to the Ganges (twice mentioned); and from the Pamirs/Himalayas southward to the Bholan (Quetta) area.

24. This situation is remarkable: if one were to apply it to a Near Eastern context, it would mean that an ancient Jerusalem temple ritual might contain Philistine, Lebanese, Ak-

kadian, Egyptian, or other "pagan" words. The Indian situation also differs remarkably from that of the Hittite empire, where the preceding non-IE language, Hattic, was actually used as the ritual language.

25. Modern recitation depends on these written texts and cannot be used in the same way as Vedic recitation.

26. Humbach 1991.

27. Discussion by Skjærvø 1995. However, the Y.Avest. local name of Bactria (*Bāxδī*) is attested earlier, in the *Atharvaveda*; see Witzel 1980. Current estimates for Zoroaster range from the fourteenth to the seventh c. B.C.E. However, an early date is indicated by the linguistic development of the name of Ahuramazdā: O.Avest. *mazdā ahura* (or *ahura mazdā*) > Y. Avest. *ahura mazdā* > Old Persian (519 B.C.E.) *A^[h]uramazdā*, already as one word. For the transfer of Zoroastrianism into Persis (the modern province of Fars, i.e., southwestern Iran), see Hoffmann 1992.

28. Note the ongoing debate, since the mid-nineteenth century, especially by S. Lévy, Przylusky, and Kuiper; see the relevant summaries in Mayrhofer 1956–1976 (*KEWA*), 1986–1996 (*EWA*); last update in Witzel 1999b, c.

29. Explained, since H. Krahe, as an "old European" layer of IE (summary in W. P. Schmid 1995); *pace* Schmid, this layer of river names has several elements that seem to differ from and to predate Proto-Indo-European; note also that many Germanic words or names in the North Sea / Baltic area belong to a pre-IE substrate. See Polomé 1990; Huld 1990.

30. Summary by Th. Lindner 1995.

31. See Hamp 1998, 328; Huld 1990.

32. For details see the author's longer article in *Sino-Platonic Papers* 129 and the study of North Central Asian religious terms (e.g., "Asura") that overlap with those in Uralic/Ket and in Indo-Iranian (Witzel 2004).

33. By 2500 B.C.E. there are camels (figurines) drawing vehicles in south Turkmenistan.

34. They could, however, have learned about camels from Central Asian hunters as represented by the Kelteminar culture.

35. A genetic link between Elamite and Dravidian has not been established; see the extensive discussion in *Current Anthropology* (McAlpin et al. 1975).

36. An account of the insurrection of Nāram-Sīn (text from the end of the third millennium B.C.E.) mentions the rulers allied against Akkade: "the man of Meluh_{ha}, the man of Aratta, the king of Marhaši . . . [another country: gap in tablet], the king of all of Elam" (Vallat 1985, 53); note also the list by the Sargonic king Rimus, who conquered Parah_{sum}, Zahar, Elam, [Ba]sin(?), and Meluh_{ha} (see Gelb 1997, 594).

37. Perhaps one of the Elamite capitals.

38. Later taken over by the Tukriš people; Steinkeller, however, locates Tukriš in the Elburz Mountains north of Tehran (1982, 265). The word "Tukriš" has been compared with "Tuyrān," "Tuyrastān," "Tokharoi," and so on (Henning 1978), and with a possible Mitanni-Indo-Aryan name "Tugra"; see *EWA* 1: 651, s.v. *tugra*. Some *RV* passages involving *tugr-* may provide a link with this Central Iranian area as well.

39. For various Elamite sources see Vallat 1985. However, Steinkeller tentatively locates it a little farther west, "in the perimeter of Kerman and Eastern Fars" (1982, 255, map p. 265). The original name of *Maharši* seems to have been **m^hparahši*, represented in Akkadian as *Parah_{sum}* (ibid., 237–238). Cf. the modern *Maškai* and other rivers in Baluchistan beginning with *Maš-*.

40. Steinkeller 1982, 1989 (see also Hiebert 1998, 147). Steinkeller connects the name of Marḥaši/Parahšum/Paraši with the name of the Persians; for this, however, note the *Parsu* reference in *RV* and *Baudhāyana Śrautasūtra* (Witzel 1999a) and a possible connection with the name of the Pashto (Morgenstierne 1927). For further Eastern Iranian / Vedic correspondences, see Witzel 1989, ch. 10, and for some Saka-like characteristics of Old Persian, see the author's longer article in *Sino-Platonic Papers* 129.

41. See Possehl 1996.

42. Which speaks against Dravidian as the language of the Indus Civilization; see further Witzel 1999b, 1999c, 2001 for other possibilities.

43. The loan word links between Sumerian and Dravidian are in need of further investigation (see Blažek and Boisson 1992 and Blažek 1992, 2002). As for the connection between the two areas, a *direct* southern route, from Sumeria via Bampur to Meluhha (provided it was Dravidian-speaking!), is not likely, as it is blocked by a large area of Elamite (from Susa up to Tepe Yahya) and by the separate linguistic area of Simaški (Bampur and Jiroft). Following the observations made in Witzel 1999b, 1999c, I suggest that that early Dravidian may have received its agricultural terminology via Aratta (Seistan) or even directly by maritime contacts.

44. Possehl 1997, with a date of 2500 – 1900 B.C.E. for the Kulli/Quetta phases.

45. Hammurabi inscription; see Henning 1978, 220.

46. Needless to say, this fragmentary information will have to be compared with the Central Asian materials presented here. At first sight, I have not detected correspondences beyond the words for “jewel” (Mit. *mani-nnu*, Ved. *maṇi*; see below), “soma” (Mit. *sauma*; *EWA* 2: 749), and the Central Iranian name “Tukriš,” possibly retained by Mit. IA **Tugra*; see above n. 38.

47. Some arguments are presented in *Sino-Platonic Papers* 129, Sections 3 – 5. For Margiana see also Erdosy 1998.

48. Note that the central Afghan highlands were occupied by Mongolians (Hazara) after the expansion of the Mongol empire — wiping out possible traces of older populations; for some details on the Central Highlands, see Witzel 2000; cf. the similar situation in the Tian Shan / Pamirs, where Mount Muzh (Tagh Ata) of our maps is derived from the old, non-IIr. name Ved. *Mūja-*, Avest. *Muža*, but was interpreted as a Turkic name. The name of the impressive, Kailash-like 24,767 feet / 8.120 meters high Muzh Tagh Ata means “ice mountain father” in Kyrgyz. Another Muzh Tagh is found northeast of Skardu in northernmost Kashmir.

49. I cannot enter into a detailed discussion of Lubotsky's recent essay (2001). His new observations, based on the Indo-Iranian words listed in *EWA*, include (1) unusual IIr. suffixes, including *-ka*, *-pa*, *-sa*, which are directly attached to the root: **stuka/stūpa* “tuft,” **pāpa* “evil,” **kac'yapa* “tortoise”; note also **jharm(i)ya* “firm structure, house”; and (2) the unusual structure of trisyllabic words with long middle syllable (CəCəCə), such as **yaviyā* “irrigation canal,” **kapāra* “vessel, dish,” **piyūša* “beestings” (first milk of a cow after it has given birth), **warāj'ha* “wild boar” (but note the early loan into FU: F-Volg. **oraše*, Finn. *oras*; Rédei 1986, 54). He had originally assumed (as per van Driem 1999) that some of these words could have only been taken over inside the subcontinent and that the language north of the Hindukush and in the Panjab was the same as that in the BMAC (or, as he now formulates, “intimately related”; Lubotsky 2001, 306; however, see Witzel 1999b, 1999c, and forthc.). This close connection is, however, unlikely and does not account for loan words with the suffix *-āša*, *-āc'a*, *-āsa* that also occur with short medial vowel (of the “Para-Munda” type; Witzel 1999b, 1999c). Proof for a CəCəCə type Panjab/BMAC language could only come from words typical for

that climate, flora, fauna, or culture. Indeed, there are at least two such words that may be indicative, *mayūra* "peacock" and *śārdūla* "tiger"; however, they have complicated etymologies (Witzel 1999b, 1999c). Only a study of Lubotsky's new words from Indo-Iranian, as found in the various strata of Vedic and Iranian texts, would allow us to determine what is securely reconstructable for Indo-Iranian, or just for the subsequent Old Indo-Aryan and Old Indo-Iranian levels. The earliest attestation of a particular "foreign" word, such as in the *RV* or the older parts of Avestan, is the only secure basis for analysis, since words from the (unknown, complex) popular local Indian level constantly emerged in the sacred speech of High Vedic and the educated speech of the Brahmins (cf. Kuiper 1991; Witzel 1989), even at much later periods (cf. IE **perd*, attested only in post-Vedic *pardati* "to fart"!). Such words could accidentally have the desired shape, CəCəCə. Note also that Lubotsky so far neglects the typical substrate words found only in Iranian, such as *kaufa* "hump, mountain" or *tūtuk* "clay" (see n. 71 and the author's longer essay in *Sino-Platonic Papers* 129), which can serve as a corrective. Lubotsky's new data, however, support my initial analysis of Central Asian loans (Witzel 1995, 1999a, 1999b, 1999c), though he does not localize this substrate beyond pointing to the towns of the BMAC and the Panjab (Lubotsky 2001, 306, echoed in van Driem 1999; for a discussion and refutation, see Witzel forthc.).

50. For instance *khara*, *rāsabha*, *gardabha*, and so on; see Rau 1980 – 1981. The Vedic word for "hemione, onager, half-ass" has not been identified yet; modern *khōr* "hemione" (still found in the Rann of Cutch), is derived from Vedic *khara* "donkey."

51. Both Vedic and Tocharian with the common Indian animal suffix *-bha* (as in *gardabha*, *ṛṣa-bha*, and *śara-bha*, for which cf. Kalash *śāra* "markhor, mountain goat"), which would point to a slightly different source.

52. The overlap of Dravidian *kalu-* and Vedic *garda-bha* "donkey" would be one of the few links of the Central Asian (BMAC?) language(s) with Dravidian (cf. McAlpin et al. 1975) and Indo-Aryan. Indo-Aryan words for domesticated animals are entirely different from Dravidian. However, Vedic *garda-bha* (EWA 1: 473), Dravidian *kalu-tai* (Burrow and Emeneau 1984, 1364) could, or rather should, be independent loans from a language of Greater Iran or Central Asia.

53. For example, in Sintashta there are only reinforced palisades with clay inside; however, the area just north of the Caucasus reportedly has clay bricks; see Witzel 2000.

54. Note that the speakers of Indo-Iranian could not have entered or been in close contact with the BMAC area earlier than the introduction of donkeys (while they could have learned about camels from Kelteminar hunters). For this passing through of Indo-Aryan speakers, note the Mitanni IA words *mani-nnu*, *Sauma*, *Tukriš*, see notes 38, 46.

55. Such as "chariot driver," "chariot warrior" (bow shooter), "spoke," "felloe," and so on; (pre-)chariots (**ratha*, not **anas* "wagon") are first found in the northern steppes at Sintashta and also in the Near East about 2000 B.C.E.; see Witzel 2000.

56. See the recent summary of Ofer Bar Yosef's work, at <http://www.harvard-magazine.com/on-line/09016.html> and now in Yasuda 2002.

57. The unfamiliar **gant-um/gand-um* > Iranian **gantum*, Indian **godum* (OIA and Drav.) was analyzed as IA *go-dhūma* "cow smoke" (cf. also Burrow and Emeneau 1984, 2226: Konda, etc., *goyi* "smoke," thus *godhūma*, an original joke form?); similarly, though differently, in Dravidian (see below n. 63).

58. For the "suffix," cf. Neo-Elam. *umi* "to grind (grain)" and PDrav. **um* "husk, chaff"; the compound "Pan-Iranian" *gant + um* (note Berger's Bur. **yund-um*) may therefore have

originally meant “wheat grain.” (Cf. also the link between PDrav. **var* “seed, grain,” and Elam. *bar* “seed”; Southworth 1988: 659–660.)

59. Bur. *guriñ*, *gureñ* (pl.), *-yárum* < **yor-um* < ***yund-* (Berger), ***yund-um*; cf. also Bur. *gur* “barley, wheat colored,” *bur* “buckwheat” (Berger 1959, 43).

60. On the interchange of *r/n*, see section 5 of the longer version of this essay in *Sino-Platonic Papers* 129.

61. Avest. *gañtuma*, M.P., N.P. *gandum*, Shughni *žindam*, Pashto *yanəm* < **gandūma?*, Khot. Saka *ganama* < **gandama* (Berger 1959, 40–41; *EWA* 2: 498); however, Brahui *xōlum* < IA **yolum* (Turner 1966, 4287).

62. As seen in Nur. *gūm*, Hindi *gohū/gehū/gahū*, and so forth. The northern form, based on Pre-Iranian **gantum*, would have resulted in Vedic **gan-dhūma* or perhaps **gandha-dhūma* “perfume smell”; cf. Turner 1966, 4020, Skt. (lex.) *gandhālu* “fragrant rice,” Pashai *gandār* “a kind of grain.” The actual Vedic form *go-dhūma* must be due to local influence by the Southern (Meluhhan) **godi* (see above on Konda *goyi*) on a Northern **gantum/gandum*; it may be due to Dravidian influence on the Panjab in the Middle/Late Rgvedic period, as *godhūma* appears only in early post-*RV* texts (Witzel 1999b, 1999c).

63. There is archeological evidence now: while Renfrew had maintained that wheat was introduced by the speakers of (hypothetical) Elamo-Dravidian, reaching India as early as 6000 B.C.E., Dorian Fuller (University College, London) now shows that wheat and barley arrived in southern India only at c. 2200 B.C.E., after having stalled for three thousand years in north-west India before farmers developed a monsoon-tolerant variety; see *Science Magazine* 294 (November 2, 2001): 989, <http://www.sciencemag.org/cgi/content/full/294/5544/988>.

64. The unfamiliar (Iranian) **godum*, adopted into India, has been reconstructed for Dravidian by Southworth (1988, 658, 660) for c. 1000 B.C.E. as **kō-tumpai* “low red plant” (perhaps because of Burrow and Emeneau 1984, 3334, Tam. *tumpai*, etc., “nettle, weed”?). The development from **tumpai* > *di*, however, is not clear. At the supposed date of 1000 B.C.E., **kōtumpai* could even be based on *RV godhūma*! This late date (along with its speculative glottochronological basis), anyhow, is now doubtful, as wheat is found in southern India by 2200 B.C.E. (see previous note). The various Elamite, Sumerian, and other loans into Dravidian will have to be compared and must be reconciled with Blažek 2002.

65. See Dani and Masson 1992, 116, 222, and so on; note, later on, also the extensive network of underground channels in Afghanistan, *karēz*, and so on.

66. Adams 1998, 376, quoting F. Hiebert (1995; cf. 1998, 231); cf., however, Francfort 1999, 451.

67. It is interesting that the Old Persian word is close to Vedic; this suggests, like several other (grammatical) items, a previous closer relationship of the two languages in northeastern Iran / Central Asia — *yav(i)yā* may have been influenced by PIIr. **naHwīya* “boatable” (from **naHu* “boat”) > Avest. (*āfš*) *nāuuaiiā* “water channel,” Skt. *nāvyā* “navigable river” (cf. Adams 1998, 373). Adams thinks that Tocharian borrowed key words for irrigation from Eastern Iranian sources (which he dates much too early, Francfort 1999, 451), though the Tocharians knew terms of primitive agriculture (plow, draft ox, harness, grain/wheat, harvest, to plow) before; see Shishlina and Hiebert 1998, 231.

68. Note also *kubja*, *kubhra* “crooked,” and N.P. *kūž*, *kūz*, Sogd. *kwzz* “humpbacked,” Khot. *kūysa* “bent,” with unclear etymology; cf. Skt. *kunṭha* “defective” (Turner 1966, 3260, 3290) ~ Iranian: N.P. *kund*, Bal. *kunt*, and the many words in Turner 1966 meaning “defective”; cf. also the river name *Kubhā* (Kabul River).

69. The Persians moved into the Persis and Anšan from northwestern Iran only after c. 700 B.C.E.; see summary in Skjærvø 1995.

70. Lead is found in the Kandahar area; see maps in Kenoyer 1995.

71. Further, a large number of verbs can be added, such as *kan "to find pleasure, please," *kram "to stride," *kroc "to shout," *kc'ā "to regard, look," *kšad "to serve food," *kšam "to be patient," and so on; or the culturally important *kaiš "to apportion magically, to teach," *mark/marc "to damage," *bhiš "to heal," with its strange derivative *bhiš-aj "healer"; or nouns such as *kaufa "mountain, hump (of a camel)"; see also n. 49.

72. Also found in Southeast Asia: Malay *sawi*, *sasawi*, Austro-Asiatic *sapi, sV(r)-sapi; cf. further the similar word, *EWA* 2: 712, 727: *śimśāpā* RV+ "Dalbergia sissoo," N.P. *śiśam*, Pashto *šawa* < *śiśampā (Turner 1966, 12424), Elam. *še-iš-šā-ba-ut* = [*šeššap*]. For the change of r/n, see n. 75 below.

73. Note that the use of *pard (*prḍāku*) has been narrowed down in India to "(spotted) snake," while the old word for "panther" has been replaced early on by *dvīpin* "the one having islands (spots)" AV+ (see *KEWA* 2: 87; *EWA* 2: 769); probably *prḍāku* has been influenced in form (and meaning: "reptile") by the local Indian *sṛdāku*/gu MS, *sṛdāku* "lizard," *sṛdara*, "snake."

74. For the sound changes involved, cf. Avest. *paridaēza* (source for our "paradise") > N.P. *pālēz* "garden."

75. On the interchange of r and n, see section 5 of the longer version of this essay in *Sino-Platonic Papers* 129; *par/nd was replaced in India early on by *dvīpin*; see above n. 73. Note that Steinkeller considers the "speckled 'dog' of Meluhha" (*ur gūn-a*, not "red dog" as often quoted!), a diplomatic gift from Marhaši, to be a leopard (*Panthera pardus*) (1982, 253). One wonders, however, why one would send such a common animal as the leopard to Ibbi-Sin, the last king of the Third Dynasty of Ur.

76. See Falk 1997 with relevant literature. Note that rhinoceros were still hunted by the Moghuls in the Peshawar area.

77. The tiger (along with the rhinoceros?), will have occurred in the swamps of Seistan. The last Central Asian tiger was reportedly shot on the Aral Lake in the seventies (cf. Masson 1992, 39), and the last one in Afghanistan, on the Amu Darya islands near Kunduz, in 1970 (Bucherer-Dietschli 1986, 95). However, the tiger (*bebr*, Horn 1893, 42; cf. Ved. *vyāghra*, *EWA* 2: 593, with folk etymology < C. Asian ***bagr*?) is still found in Iran, in the Elburz and Kopet Dagh mountains. Its absence in the RV, in contrast to its prominence in the seals of the Indus Civilization (along with the R̥gvedic absence of the leopard that has been found depicted from the early Turkmenistan pottery onward), may be because it did not belong to the traditional imagery of the Indo-Aryans, who preferred the lion (*simha*). A possible reason may be the preponderance of lion images in religion (and the absence of panther and tiger), a religious choice (Francfort 1994) already made by the people of the BMAC area, through which the speakers of Old Indo-Aryan must have passed. A new study of early Indo-Iranian animal designations and their respective importance in religion (cf. W. Voigt 1937) is a desideratum. In view of these uncertainties, it remains entirely doubtful whether the languages of the BMAC and of Harappa were related at all, as Lubotsky (2001, 306) believes (see above, n. 49).

78. First collected and discussed, as far as is represented in *EWA*, by Lubotsky 2001; now by G. Thompson, at the Fifth Harvard Round Table, May 2003: <http://www.fas.harvard.edu/~sanskrit/RoundTableSchedule03.html>.

79. The last three words may, however, be late loans into *Vidēvdād*, from Old Indo-Aryan in the late first millennium B.C.E.; but note Pinault on *c'arwa*, n. 88.

80. Indo-Iranian **sauma*, Ved. *soma*, O.P. *hauma*, Avest. *haoma*, from the root *su* "to press," thus **sauma* "the pressed drink"; note also the name of the Saka Haumavarga, "the Soma twisting/pressing Sakas," in Central Asia, attested in Old Persian inscriptions; see discussion in Witzel 2004.

81. See Parpola 1995, Nyberg 1995 (= ephedra), and the recent Leiden workshop on the problem; its results have partially been published in *Electronic Journal of Vedic Studies* 9 (May 2003), ed. J. Houben. For ephedra in the Kalash valleys of Chitral, see next note.

82. Thompson 2001 has also appeared in the *Electronic Journal of Vedic Studies* 9. Note that the word for *Asafoetida* (a plant producing resin) is similar in shape: Ilr. *anc'u* could go back to an older pre-Ilr. form **ank'u* that is found in Toch. as *ankwaš*, Chin. *yāngkuì* (a resin of *Asafoetida sinica*, a low-grade stimulant); see Pulleyblank 1962. However, Lubotsky simply subsumes it under *Wanderwörter* of unknown etymology (1998, 379). Pinault (2003) further connects the Tocharian word for "iron," A **ańcu* (adj. *ańcuwaši* "made of iron"), B *eńcuwo* (adj. *eńcuwanne*), taken from the rusty color of iron ore, just as the soma plant and its juice are called "rusty brown" (as is the powder of some ephedra varieties; cf. below). Interestingly, Arrianos, *Anabasis* iii, ch. 28, is aware of it as *silphion*, growing in the "Caucasus" (Hindukush); this is identified as *Asafoetida*, liked by sheep; see Wirth and Hinüber 1985, 887, n. 111. Note that the same is said about ephedra = Kalash *sámani* (< **sumanas*; Turner 1966, 13492): it grows in the higher mountain meadows, is eaten by goats, and is used as an ingredient in making chewing tobacco (*nazwár*); it is, according to Morgenstierne, who went to the Kalash Valleys in 1929, a "small shrub resembling heather, with stiff, straight twigs, growing in the high mountain valleys." The latter description is the same as that given for *hum* in Afghanistan: "A stiff bush, three feet high, with closely arranged, vertical, leafless, articulated stems. It grows on stony, infertile ground. The fruits are red and fleshy and are eaten by children. The stems are used for dying yellow, and are used, as powder, for chewing [tobacco] and as snuff. The Afridi tribes crush the stems and soak them in cold water. The extract derived from it is used as medicine against fever. The plant grows in all of Turkestan, north and middle Persia, in northern and eastern Afghanistan and in the northwestern Himalayas" (K. F. Geldner, *Der Rig-Veda*, Harvard Oriental Series, 35 [Cambridge, 1951], introduction to vol. 3, 2).

83. See Staal (2001, 2003) with a localization; the *RV* speaks of the best variety, Maujavata, from the mountain Mūjavant (*AV*), cf. Avestan *Muža*, and note even now the mountain (range) Muzh Tagh Ata, and the river Muzh-Kol, on the Xinjiang/Tajikistan border (as well as north of Skardu); see Witzel 2000 and n. 48 above.

84. See Witzel 2000 for a discussion of non-Indo-Iranian Hindukush and Pamir mountain peoples.

85. Sarianidi 1992: 34; this custom continued down even to the Greek period in Bactria (Ai Khanum).

86. In detail: **atharwan*, Ved. *atharvan* "ancient priest, sorcerer," Avest. *āθrauuān* "priest" (*EWA* 1: 60); Pinault 2003 connects this with Toch. B *etre*, A *atār* "hero"; **r̥ši* "seer, poet," Ved. *r̥ši*, cf. O.Avest. *ərəšiš* "(person) full of enthusiasm for god," Yasna 31.5, *ərəšišīia* (*EWA* 1: 261); **uc'ig* "sacrificing priest," Ved. *uśij*, *auśija*, Avest. *usij* "sacrificer following non-Zoroastrian ritual/belief" (*EWA* 1: 234); **yātu* "black magic," Ved. *yātu*, Avest. *yātu* "sorcery, sorcerer," N.P. *jādū* "sorcerer." Note also **kaiš* "to apportion magically, to teach" (cf. *EWA* 1: 359).

87. The Vedic/Avestan lists have (1) Ved. *Varuṇa*, Mitanni *Uruna* (substituted by Avest. *Abura Mazdā*?; with a local(?) *-una* suffix otherwise rarely found in IE, IIr. (in *Yamunā*, *taruṇa*, *aruṇa*; cf. Avest. *tauruna*, *auruna*), (2) Ved. *Mitra*, Mit. *Mitra*, Avest. *Miθra*, "god Agreement," (3) *Arya-man*/*Aⁱriia-man* ("Arya-hood," god of marriage exchange, an artificial formation from an adjective with the deverbal derivative suffix *-man*, similarly Avest. *X^vaē-tu* "selfhood" > "belonging to a family"), (4) *Bhaga/baya* (O.P. *baga*) "god Share," (5) *Aṃśa* "god Lot," (6) *Dhātṛ/dātar*- "god Apportioner/Creator," *Dakṣa* "god Cleverness," *Mārtāṇḍa Vivasvant* / *Gaiiō marətan*, *Viuuanjhuuant*. These new deities are active in social relationships, respectively, in the world, tribe, clan (marriage), family, and individual; Vivasvant is the ancestor of Manu and Yama/Yima and, thus, of human beings (or at least of all *ārya* lineages). They are not found in Indo-European (Slav. *bog* "god" is a loan from Northern Iranian *baga* "god"). Their Indo-Iranian origin may be attributable to socioreligious conditions of the Northern Steppes (water rights, oral agreements, and so on), as is indicated by the several correspondences of these words in Uralic and Yeneseian (discussion in Witzel 2004). However, the BMAC population had an altogether different local religion, with a preponderance of an (agricultural) fertility goddess and her antipode, a polymorphous male scaled-skin dragon deity (combining several animal and human forms and allied with snakes and lions), who is fought by an eagle-headed hero (Francfort 1994; see also Witzel 2004).

88. Details: **c'arwa* "name of a god," Ved. *Śarva* "name of Rudra," Iran. *Sauruua* "a demon"; Pinault 2003 connects the Tocharian word for "hunter": Common Toch. **śerwæ*, B *śer(u)we*, A *śaru* "hunter" and N. Iran. words: Osset. Digor *sorun*, *surd*, Iron *sūryn*, *syrd* "to hunt," Khotan. *ha-sūrā*, *hu-sure* "quarry, hunted beast"; **indra* "name of the 'king' of the present generation of the gods," Ved. *Indra*, Mitanni *Indara*, Avest. *Indara*, "a demon"; **g(h)andharw/b(h)a*, Ved. *Gandharva* "name of a semidivine spirit of lust and procreation," Avest. *gaṇdarəβa* "a demon, monster."

89. Originally, Lubotsky had apparently thought that this group of words belonged to a still older layer of loans in Indo-Iranian.

90. Doubtful, at least for **c'arwa*, unless one assumes a root **c-*, as can happen in Caucasian languages; cf. Northeast Caucasian river names in the author's longer essay in *Sino-Platonic Papers* 129. See, however, n. 88 for a Tocharian reflex.

91. For details, see above notes 86–87, for example, *Arya-man*. Is the Indo-Iranian suffix *-man/van* replacing a BMAC suffix *-arwa*, whose use would have resulted in the difficult form Ved. **arya-rva/aryārva*? Cf., finally, the rare *-vala* suffixes in Vedic (*RV a-kṛṣi-vala*).

92. For an initial investigation, see Witzel 2004.

93. For a beginning, see Witzel 2004.

REFERENCES

- Adams, D. Q. 1998. On the History and Significance of Some Tocharian B Agricultural Terms. In Mair 1998, 1: 372–378.
- Anttila, R. 1989. *Historical and Comparative Linguistics*. Amsterdam and Philadelphia: John Benjamins.
- Balkan, K. 1954. *Kassitenstudien I: Die Sprache der Kassiten*. New Haven: American Oriental Society.
- Bengtson, J. D. 1999. A Comparison of Basque and (North) Caucasian Basic Vocabulary. *Mother Tongue* 5: 41–57.

- . 2001. Genetic and Cultural Linguistic Links between Burushaski and the Caucasian Languages and Basque. Paper at the Third Harvard Round Table on the Ethnogenesis of Central and South Asia, May 2001. (<http://www.fas.harvard.edu/~sanskrit/RoundTableSchedule01.html>).
- Berger, H. 1959. Die Burušaski-Lehnwörter in der Zigeunersprache. *Indo-Iranian Journal* 3: 17–43.
- Blažek, V. 1992. The New Dravidian-Afroasiatic Parallels. In *Nostratic, Sino-Caucasian, Austric and Amerind*, ed. V. Shevoroshkin, 150–165. Bochum: Brockmayer.
- . 2002. Elam: A Bridge between the Ancient Near East and Dravidian India? *Mother Tongue* 7: 123–143.
- Blažek, V., and C. Boisson. 1992. The Diffusion of Agricultural Terms from Mesopotamia. *Archiv Orientalní* 60: 16–37.
- Bondaruk, G. P., È. M. Murzaev, V. A. Nikonov, and V. V. Tsybul'ski, eds. 1964. *Toponimika vostoka: Novye issledovaniia*. Moscow: Nauka.
- Bucherer-Dietschli, P., and C. Jentsch. 1986. *Afghanistan*. Ländermonographie, bearbeitet von der Arbeitsgemeinschaft Afghanistan. Liestal: Stiftung Bibliotheca Afghanica.
- Burrow, T., and M. B. Emeneau. 1984. *A Dravidian Etymological Dictionary*. Second ed. Oxford: Clarendon Press.
- Colarusso, J. 1998. Languages of the Dead. In V. Mair 1998, 431–447.
- . 2002. Remarks on the Anau and Niyā Seals. *Sino-Platonic Papers* 124 (August).
- Dani, A. H., and V. M. Masson, eds. 1992. *History of Civilisations of Central Asia*, Vol. 1: *The Dawn of Civilisation: Earliest Times to 700 B.C.E.* Paris: UNESCO Publishing.
- Eichler, E., G. Hilty, H. Löffler, H. Stegler, and L. Zgusta, eds. 1995. *Namenforschung. Name Studies. Les nomes propres: Ein internationales Handbuch zur Onomastik. An International Handbook of Onomastics. Manuel international d'onomastique*. Vol. 1. Berlin: de Gruyter.
- Eilers, W. 1982. *Geographische Namensgebung in und um Iran: Ein Überblick in Beispielen*. Bayerische Akademie der Wissenschaften, Phil.-Hist. Klasse, Sitzungsberichte, 5. Munich.
- . 1987. *Iranische Ortsnamenstudien*. Sitzungsberichte der Österreichischen Akademie der Wissenschaften, Phil.-Hist. Klasse, 465. Vienna.
- Erdosy, G., ed. 1995. The Indo-Aryans of Ancient South Asia. *Indian Philology and South Asian Studies*, ed. A. Wezler and M. Witzel, vol. 1. Berlin and New York: de Gruyter.
- . 1998. Language, Ethnicity, and Migration in Protohistoric Margiana. In *The Archaeological Map of the Murghab Delta: Preliminary Report 1990–1995*, ed. A. Gubaev, G. Kosholenko, and M. Tosi, 141–147. Rome: IsAIO.
- Falk, H. 1997. The Purpose of R̥gvedic Ritual. In *Inside the Texts, beyond the Texts: New Approaches to the Study of the Vedas*, ed. M. Witzel, 67–88. Cambridge: Harvard Oriental Series, Opera Minora.
- Francfort, H.-P. 1989. *Fouilles de Shortughai: Recherches sur l'Asie centrale protohistorique, par Henri-Paul Francfort avec des contributions de Ch. Boisset [et al.]*. Mémoires de la Mission archéologique française en Asie centrale. Paris: Mission archéologique française en Asie centrale: de Boccard.
- . 1990. *Nomades et sédentaires en Asie centrale: Rapports de l'archéologie et de l'ethnologie: actes du colloque franco-soviétique, Alma Ata (Kazakhstan), 17–26 octobre 1987*. Paris: Éditions du Centre national de la recherche scientifique, distributed by Presses du CNRS.

- . 1994. The Central Asian Dimension of the Symbolic System in Bactria and Margiana. *Antiquity* 68: 406–418.
- . 1999. Compte rendu: V. H. Mair, 1998. *Bulletin des études françaises de l'Extrême Orient* 86: 449–458.
- Francfort, H.-P., and E. M. Kuz'mina. 1998. Du nouveau dans la chronologie de l'Asie Centrale du Chalcolithique à l'Age du Fer: Rapport du Groupe de travail "Asie." In *Actes du colloque "C14 Archéologie,"* 1998, 467–469. Paris.
- Gelb, I. J. 1997. Makkan and Meluhha. *Early Mesopotamian Sources, Revue d'Assyriologie* 64 (1970): 1–7. Also in *A Source-Book of Indian Archaeology*, ed. F. R. Allchin and D. Chakrabarti, vol. 2: *Settlements, Technology and Trade*. Delhi: Munshiram.
- Gryunberg, A. L. 1980. Zametki po toponimii afghanskogo Badakhshana. In Murzaev 1980, 165–169.
- Hamp, E. 1998. Whose Were the Tocharians? In V. Mair 1998, 307–346.
- Henning, W. 1978. The First Indo-Europeans in History. In *Society and History: Essays in Honour of Karl August Wittvogel*, ed. G. L. Ulmen, 215–230. The Hague–Paris–New York: Mouton.
- Hiebert, F. T. 1995. South Asia from a Central Asian Perspective. In G. Erdosy 1995, 192–212.
- . 1998. Central Asians on the Iranian Plateau: A Model for Indo-Iranian Expansionism. In V. Mair 1998, 148–161.
- . 2001. The Recently Discovered Bronze Age Inscription (2300 BC) from Anau, Central Asia. Paper at Third Harvard Round Table on the Ethnogenesis of South and Central Asia, May 12–14 (<http://www.fas.harvard.edu/~sanskrit/RoundTableSchedule.html>). 2002.
- . The Context of the Anau Seal. *Sino-Platonic Papers* 24 (August).
- Hiebert, F. T., and C. C. Lamberg-Karlovsky. 1992. Central Asia and the Indo-Iranian Borderlands. *Iran* 1992: 1–15.
- Hoffmann, K. 1992. *Aufsätze zur Indoiranistik*. Ed. S. Glauch, R. Plath, S. Ziegler. Vol. 3. Wiesbaden.
- Horn, P. 1893. *Grundriss der neupersischen Etymologie*. Strassburg: Trübner.
- Houben, J. E. M., ed. 1999/2003. Soma Workshop, Leiden, July 1999. 1. The Soma-Haoma Problem: Introductory Overview and Observations on the Discussion (J. E. M. Houben); 2. Report of the Workshop (J. E. M. Houben); 3. Report Concerning the Contents of a Ceramic Vessel Found in the "White Room" of the Gonur Temenos, Merv Oasis, Turkmenistan (C. C. Bakels); 4. Margiana and Soma-Haoma (Victor I. Sarianidi); 5. Soma and Ecstasy in the R̥gveda (G. Thompson). *Electronic Journal of Vedic Studies* 9 (2003). <http://users.primushost.com/~india/ejvs/>.
- Huld, M. E. 1990. The Linguistic Typology of the Old European Substrata in North Central Europe. *Journal of Indo-European Studies* 18: 389–411.
- Humbach, H. 1984. A Western Approach to Zarathustra. *Journal of the K. R. Cama Oriental Institute* 51: 15–31.
- . 1991. *The Gāthās of Zarathustra and the Other Old Avestan Texts*. In collaboration with J. Elfenbein and P. O. Skjærvø. Pt. 1. Heidelberg.
- Humbach, H., and Susanne Ziegler. 1998–. *Ptolemy, Geography*, book 6: *Middle East, Central and North Asia, China*. Wiesbaden: Reichert.
- Kenoyer, J. Mark. 1995. Interaction Systems, Specialised Crafts and Culture Change: The Indus Valley Tradition and the Indo-Gangetic Tradition in South Asia. In *The Indo-*

- Aryans of Ancient South Asia*, ed. G. Erdosy, 213 – 257. Berlin and New York: de Gruyter.
- Klein, J. 1997. On Verbal Accentuation in the Rigveda. In Witzel 1997, 139 – 166.
- Kuiper, F. B. J. 1948. *Proto-Munda words in Sanskrit*. Amsterdam: Noord-Hollandische Uitgevers Maatschappij.
- . 1991. *Aryans in the Rigveda*. Amsterdam and Atlanta: Rodopi.
- . 1997. *Selected Writings on Indian Linguistics and Philology*, ed. A. Lubotsky, M. S. Oort, and M. Witzel. Amsterdam and Atlanta: Rodopi.
- Lindner, Th. 1995. Griechische (ind. mykenische) Ortsnamen. In Eichler et al. 1995, 690 – 705.
- Lubotsky, A. 1998. Tocharian Loan Words in Old Chinese: Chariots, Chariot Gear, and Town Building. In Mair 1998, 379 – 390.
- . 2001. Indo-Iranian Substratum. In *Early Contacts between Uralic and Indo-European: Linguistics and Archaeological Considerations*, ed. C. Carpelan et al., 301 – 317. Mémoires de la Société Finno-Ougrienne, 242. Helsinki.
- McAlpin, D. W. 1981. *Proto-Elamian-Dravidian: The Evidence and Its Implications*. Transactions of the American Philosophical Society, 71. Philadelphia.
- McAlpin, D. W., et al. 1975. Comment on David W. McAlpin's Elamite and Dravidian: Further Evidence of Relationship. (With discussion by M. B. Emeneau, W. H. Jacobson, F. B. J. Kuiper, H. H. Paper, E. Reiner, R. Stopa, F. Vallat, R. W. Wescott, and a reply by McAlpin). *Current Anthropology* 16: 105 – 115.
- Mair, V., ed. 1998. *The Bronze Age and Early Iron Age Peoples of Eastern Central Asia*. 2 vols. Journal of Indo-European Studies Monograph 26. Washington and Philadelphia: The Institute for the Study of Man / University of Pennsylvania Museum Publications.
- . 2001. *Notes on the Anau Inscription*. *Sino-Platonic Papers* 112 (July 2001).
- Masson, V. M. 1992. The Environment. In Dani and Masson 1992, 29 – 44.
- Mayrhofer, M. 1956 – 1976. *Kurzgefasstes etymologisches Wörterbuch des Altindischen*. Heidelberg. (KEWA.)
- . 1986 – 1996. *Etymologisches Wörterbuch des Altindoarischen*. Heidelberg. (EWA.)
- Meadow, R. 1983. The Equids of Mehrgarh, Sibri and Pirak: An Osteological Evidence for the Introduction of the Horse to South Asia. Paper presented at the Seventh Conference of South Asian Archaeology, Brussels, July 4 – 8.
- Morgenstierne, G. 1927. *An Etymological Vocabulary of Pashto*. Oslo: J. Dybwad.
- Murzaev, È. M. 1980. *Onomastika vostoka*. Sbornik statei, Ordena Trudovogo Krasnogo Znameni Institut Vostokovedeniia, Institut etnografii. Moscow: Nauka.
- Nyberg, H. 1995. The Problem of the Aryans and the Soma: The Botanical Evidence. In Erdosy 1995, 382 – 406.
- Pakhalina, T. N. 1976. O proistoskhozhdenii toponimov Ishkashim, Yazguliām i Vakhān. In *Iranskoe iazykoznanie: Istoriia, etimologii, tipologii k 75-letiiu prof. V.I. Abaeva*, ed. V. S. Rastorguev. Moscow: Nauka.
- Parpola, A. 1994. *Deciphering the Indus Script*. Cambridge: Cambridge University Press.
- . 1995. The Problem of the Aryans and the Soma: Textual-Linguistic and Archaeological Evidence. In Erdosy 1995, 351 – 381.
- Pinault, G.-J. 2003. Further Links between the Indo-Iranian Substratum and the BMAC Language. Paper given at the Twelfth World Sanskrit Conference, Helsinki, July 14 – 18.
- Polomé, E. C. 1986. The Non-Indo-European Component of the Germanic Lexicon. In